

Capítulo 8

Micrófonos

8.1. Introducción

El primer elemento de la cadena de audio es el micrófono, un *transductor* capaz de convertir señal sonora en señal eléctrica. Con mayor precisión, convierte presión sonora en tensión. Estudiaremos algunos parámetros y especificaciones importantes, así como las estructuras constructivas de los micrófonos dinámicos y capacitivos.

8.2. Sensibilidad

A los fines de conectar un micrófono con el resto de los componentes, es importante conocer cuánta tensión produce ante una dada presión sonora. En el capítulo 6 vimos que ante determinada amplitud de la presión, en los terminales del micrófono se producía cierto valor de tensión. Dijimos en ese momento que no tenía sentido afirmar, por ejemplo, que "la presión es mayor que la tensión". Pero sí tiene sentido determinar la relación que hay entre la tensión y la presión. Esa relación se denomina **sensibilidad** del micrófono. En la Figura 8.1 se muestran las formas de onda de la presión y la tensión.

La sensibilidad de un micrófono puede definirse como *el cociente entre la tensión producida y la presión que le da origen*, es decir

$$S = \frac{V}{P}$$

Recordemos que la unidad que se utiliza internacionalmente para medir la presión es el pascal, abreviado Pa (que equivale aproximadamente a 10 millonésimas, es decir 1 cienmilésima, de la presión atmosférica), de modo que la sensibilidad de un micrófono se expresa en volts por pascal (V/Pa). Otra manera muy difundida de expresar la sensibilidad es en dB referidos a 1 V/Pa. En ese caso, llamando **sensibilidad de referencia**, S_{ref} , a 1 V/Pa, se obtiene con esta fórmula:

$$S|_{dB} = 20 \log_{10} \frac{S}{S_{ref}}$$

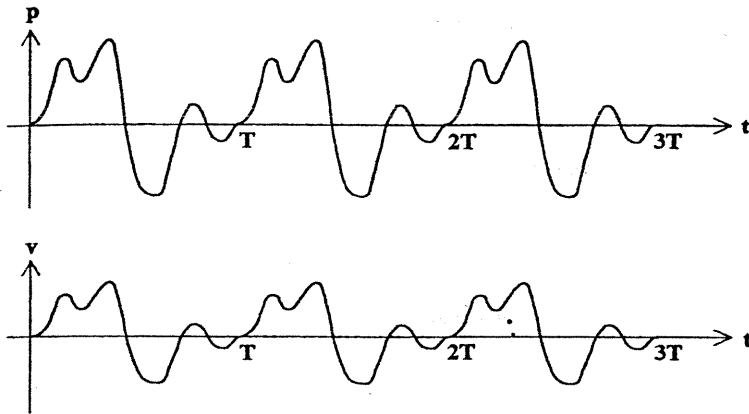


Figura 8.1. Presión sonora y su análoga, la tensión generada por un micrófono.

Por ejemplo, un micrófono que ante una presión sonora de 0,2 Pa desarrolla una tensión de 1 mV, tendrá una sensibilidad

$$S = \frac{0,001 \text{ V}}{0,2 \text{ Pa}} = 0,005 \frac{\text{V}}{\text{Pa}},$$

que en dB será

$$S_{\text{dB}} = 20 \log_{10} \frac{0,005}{1} = -46 \text{ dB}.$$

El signo (-) es porque la sensibilidad es menor que la de referencia, es decir debe calcularse el logaritmo de un número menor que 1, que es negativo.

NOTA: A veces (especialmente en especificaciones de micrófonos que llevan muchos años en el mercado), en lugar de utilizar como referencia 1 V/Pa se utiliza 1 V/ μbar . El μbar (microbar) es una unidad de presión igual a 0,1 Pa, por lo cual la sensibilidad referida a 1 V/ μbar resulta 20 dB menor (más negativa) que al referirla a 1 V/Pa. Así, en el ejemplo anterior tendríamos $S_{\text{dB ref 1 V}/\mu\text{bar}} = -66 \text{ dB}$.

Como segundo ejemplo, nos preguntamos qué tensión proporcionará este mismo micrófono ante un nivel de presión sonora de 94 dB. Sabemos que

$$94 \text{ dB} = 20 \log_{10} \frac{P}{P_{\text{ref}}},$$

donde P_{ref} es la presión de referencia, que expresada en micropascales, μPa (es decir $1/1.000.000$ de pascal), es 20 μPa . Nos interesa obtener P. Para ello, primero escribimos

$$\log_{10} \frac{P}{P_{ref}} = \frac{94}{20} = 4,7 .$$

De allí, aplicando la operación inversa del logaritmo, que consiste en elevar 10 a una potencia igual al logaritmo, se tiene

$$\frac{P}{P_{ref}} = 10^{4,7} = 50119 .$$

Entonces

$$P = 50119 \times P_{ref} = 50119 \times 20 \mu\text{Pa} \cong 1 \text{ Pa} .$$

El valor de 94 dB es un valor bastante utilizado precisamente porque corresponde a una presión sonora de 1 Pa. Con este valor, conociendo la sensibilidad, se obtiene la tensión:

$$V = S \times P = 0,005 \times 1 = 0,005 \text{ V} .$$

La señal de tensión de los micrófonos es, normalmente, muy pequeña (salvo para niveles de presión sonora muy altos), lo cual implica que está muy expuesta a los ruidos eléctricos. Por esta razón *es preciso utilizar cables y conexiones de excelente calidad para los micrófonos*, así como **preamplificadores** de bajo ruido.

Para simplificar el cálculo de la presión, la **Tabla 8.1**, reproducida del capítulo 1, da la equivalencia entre la presión y el nivel de presión sonora para varios valores.

Tabla 8.1. Conversión entre el valor de la presión y el nivel de presión sonora.

NPS (dB)	P (Pa)
120	20,0
110	6,3
105	3,6
100	2,0
95	1,1
90	0,63
85	0,36
80	0,20
75	0,11
70	0,063
60	0,020
50	0,0063
40	0,0020
30	0,00063
20	0,00020
10	0,000063
0	0,000020

8.3. Respuesta en frecuencia

Una característica importante de cualquier componente de un sistema que procesa señal, en particular de los micrófonos, es su respuesta en frecuencia. La respuesta en frecuencia de un micrófono es una gráfica que indica la sensibilidad en dB en función de la frecuencia. El aspecto típico de la respuesta en frecuencia de un micrófono se muestra en la Figura 8.2. Se observa que la respuesta *no es plana*, vale decir que no es

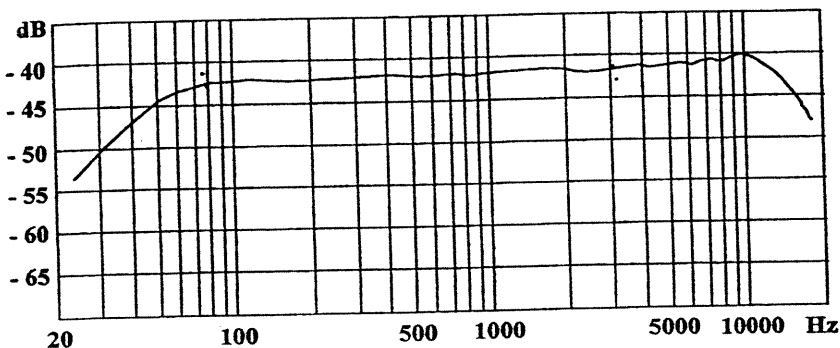


Figura 8.2. Curva de la respuesta en frecuencia de un micrófono típico.

constante con la frecuencia. Esto significa que ante dos sonidos de diferente frecuencia, por ejemplo 30 Hz y 10 kHz, pero *idéntica amplitud*, el micrófono generará *tensiones diferentes*. En este ejemplo, la sensibilidad para 30 Hz es de -50 dB, mientras que para 10 kHz es de -40 dB, lo cual hace una diferencia de 10 dB. Esto implica que la tensión generada por el micrófono a 10 kHz será (a cálculo hecho) más de 3 veces mayor que la generada a 30 Hz.

También se nota en la respuesta cierta *irregularidad* (fluctuaciones) en alta frecuencia. Esto es una consecuencia directa de que la longitud de onda a esas frecuencias ya es comparable al tamaño del micrófono (por ejemplo a 10 kHz la longitud de onda es de 3,45 cm), lo cual hace que el propio micrófono interfiera en el campo sonoro causando el equivalente de "sombras" acústicas sobre sí mismo, que dependen mucho de la longitud de onda.

Finalmente, se aprecia que existe una banda de frecuencias, que en el ejemplo abarca desde alrededor de 50 Hz hasta unos 15.000 Hz, en que la respuesta es bastante plana. Los extremos se denominan respectivamente *frecuencia inferior* y *frecuencia superior*, definidas como aquellas frecuencias por debajo de la cual y por encima de la cual la sensibilidad cae 3 dB (o en algunas especificaciones, 1 dB) por debajo del valor a 1 kHz. Cuando se desea dar una idea rápida de la respuesta en frecuencia de un micrófono, se especifican las frecuencias inferior y superior, lo cual en general es suficiente para decidir si un micrófono es o no adecuado para determinada aplicación.

8.4. Direccionalidad

Otra característica importante en los micrófonos es su direccionalidad. Debido a su construcción, y a los principios de la Acústica, la sensibilidad de un micrófono varía según el ángulo respecto a su eje desde donde viene el sonido. En la Figura 8.3 se ilustra este hecho.

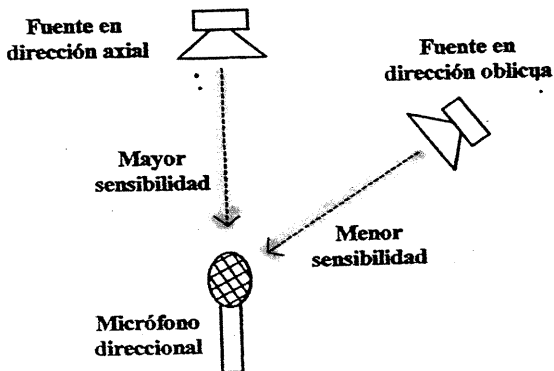


Figura 8.3. Efecto sobre la sensibilidad de un micrófono direccional (por ejemplo cardioide) de las diversas orientaciones de la fuente.

Se pueden indicar las características direccionales de un micrófono por medio de un **diagrama direccional** o **diagrama polar** como el que se muestra en la Figura 8.4.

En este tipo de diagrama se indica cómo varía de la sensibilidad del micrófono con el ángulo entre la fuente sonora y el **eje principal**, es decir aquella dirección de máxima sensibilidad. En el ejemplo de la Figura 8.4, por ejemplo, a los 90° la sensibilidad es unos 6 dB menor que en el eje principal. ⊕

El **patrón direccional** (forma del diagrama polar) de un micrófono varía con la frecuencia, debido a que para altas frecuencias, la longitud de onda es pequeña, comparable al tamaño del propio micrófono, que proyecta sobre sí mismo "sombras" acústicas que dependen de la orientación y de la longitud de onda (y por lo tanto de la frecuencia). En la Figura 8.5 se repite el diagrama polar de la Figura 8.4, incluyendo otras dos frecuencias.

Se han popularizado diversos **patrones direccionales**, cada uno destinado a un tipo específico de aplicaciones. El **patrón omnidireccional**, cuyo diagrama polar se ilustra en la Figura 8.6, tiene la misma sensibilidad en todas las direcciones, por lo cual no requiere ser apuntado hacia la fuente. Este tipo de micrófono se utiliza precisamente cuando se requiere captar sonido ambiental, sin importar su procedencia. Los micrófonos omnidireccionales, en general tienen menor variación del patrón polar con la frecuencia, razón por la cual no "colorean" el sonido proveniente de direcciones diferentes del eje principal, es decir que no presentan picos importantes en la respuesta en frecuencia.

En la Figura 8.4 se muestra un patrón cardioide. Estos micrófonos son bastante direccionales, reduciéndose muchísimo su sensibilidad en la dirección opuesta a la principal (180°). Debido a su característica direccional, los micrófonos cardioides tienen la

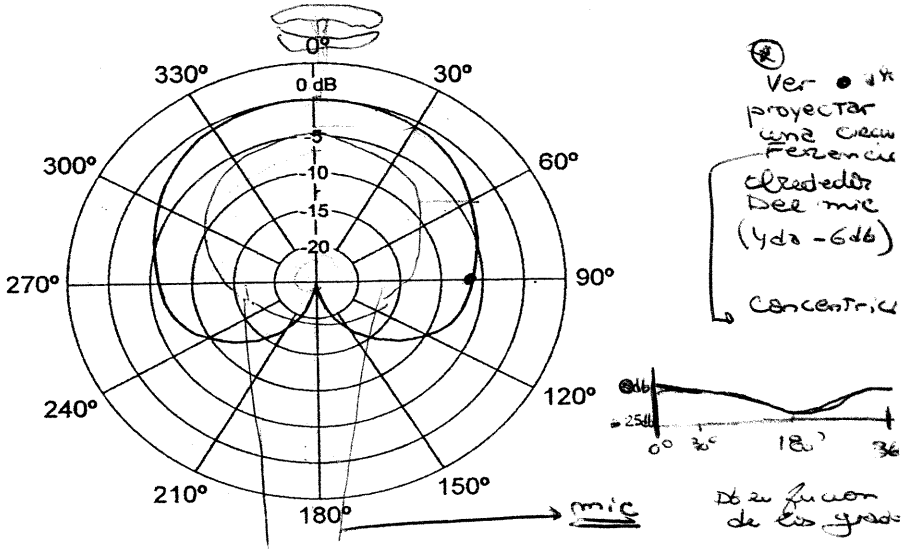


Figura 8.4. Un ejemplo de diagrama direccional o diagrama polar de micrófono. En él se indica cómo varía la sensibilidad con la dirección de procedencia del sonido, respecto a la sensibilidad máxima (0 dB), que corresponde a la dirección principal del micrófono. Este ejemplo corresponde a un micrófono cardioide.

particularidad de que cuando la fuente se aproxima mucho al micrófono (3 ó 4 cm), la respuesta en frecuencia cambia, aumentando la sensibilidad en las bajas frecuencias. Esto se denomina efecto de proximidad, y es utilizado por los vocalistas para engrosar el tono de su voz.

Una de las principales aplicaciones del patrón cardioide (también llamado direccional o unidireccional) es la de tomar sonido de una fuente determinada cuya posición es bastante estable, como por ejemplo un instrumento musical, rechazando lo más posible los sonidos provenientes de otras fuentes. Así, la captación del ruido ambiente se reducirá considerablemente, ya que el ruido es multidireccional, es decir que proviene de todas las direcciones. Un micrófono omnidireccional, lo captará en su totalidad, mientras que uno cardioide tomará sólo una parte de dicho ruido.

Dado que en estos micrófonos el patrón polar cambia bastante con la frecuencia (Figura 8.5), al captar sonidos laterales, estos resultarán "filtrados", acentuándose algunas frecuencias por sobre otras, lo cual implicará que dichos sonidos laterales estarán bastante distorsionados, o "coloreados". Los micrófonos cardioides de mejor calidad están diseñados para evitar lo más posible estas fluctuaciones con la frecuencia de sus patrones polares.

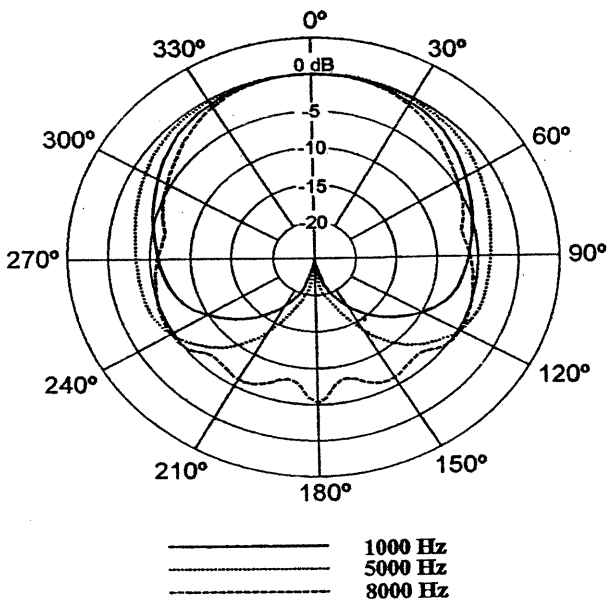


Figura 8.5. Variación con la frecuencia del diagrama polar del micrófono cardioide de la Figura 8.4. Las diferentes curvas responden al diferente patrón de "sombras" acústicas para cada longitud de onda.

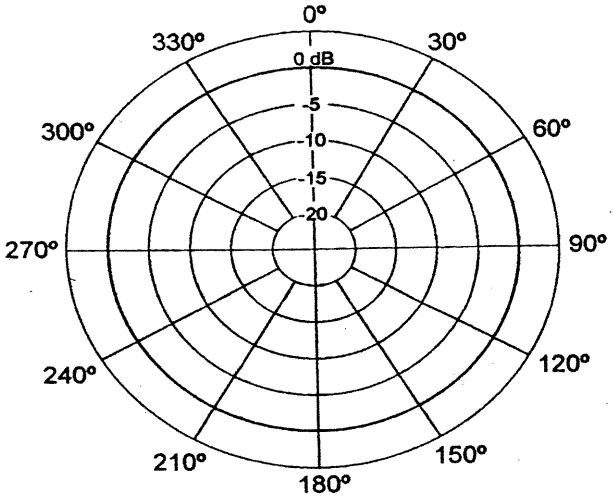


Figura 8.6. Patrón polar omnidireccional

micrófonos de los cuatros



Otro patrón polar difundido es la figura de ocho, llamada así por tener la forma de un 8 (Figura 8.7). Este tipo de micrófono podría denominarse también bidireccional, ya que es fuertemente direccional en las dos direcciones paralelas al eje principal. En la dirección perpendicular a este eje, por el contrario, la sensibilidad es nula, por lo que permite eliminar casi por completo la captación de ruidos provenientes de dichas direcciones.

Como los cardioides, exhiben también el efecto de proximidad, aumentando la sensibilidad a los graves cuando la fuente se acerca mucho al micrófono.

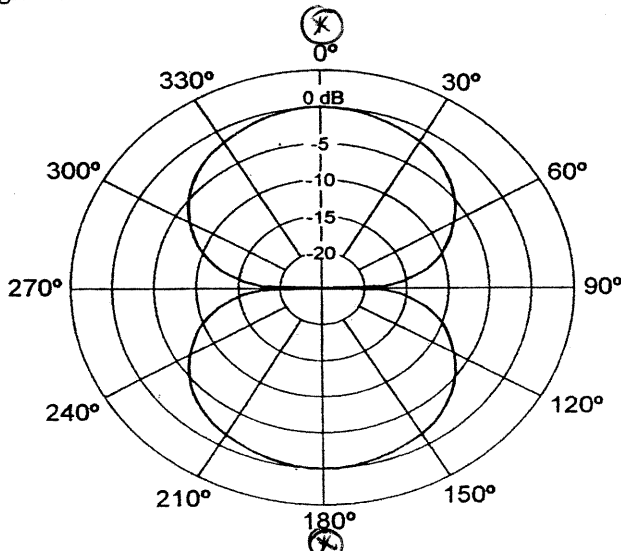


Figura 8.7. Patrón polar figura de ocho.

Dado que estos micrófonos se caracterizan por rechazar las señales acústicas provenientes de los lados de una fuente, son útiles para minimizar la captación de señal proveniente de un músico o cantante que se encuentra al lado del que se pretende tomar con el micrófono. También se utilizan para grabaciones estereofónicas (cuando se pretende crear la imagen sonora estéreo directamente desde la grabación y no por mezcla posterior), colocando para ello dos micrófonos a 90° entre sí. Esta configuración se denomina X-Y. De esta forma, la señal captada por cada micrófono será rechazada por el otro, contribuyendo a crear una mayor independencia o separación de los canales.

Además de los tipos principales descritos, existen en el mercado micrófonos con otros patrones polares, como por ejemplo el subcardioide (menos direccional que el cardioide), hipercardioide (similar al cardioide pero con un ángulo de captación todavía menor, a costa de la existencia de un pequeño lóbulo en la dirección opuesta a la principal), o el lobular (muy direccional, con un lóbulo que abarca ángulos de captación tan cerrados como 90°). La aplicación de estos micrófonos es bastante específica, y conviene en cada caso aplicarlos según las indicaciones del fabricante.

En general, los micrófonos direccionales (cardioides, figura de ocho, hipercardioides, etc.) tienen peor respuesta en frecuencia que los omnidireccionales. Esto se debe a

que según se mostraba en la Figura 8.5, el patrón polar de los direccionales varía con la frecuencia, y por lo tanto para direcciones diferentes de la principal la respuesta en frecuencia tiene fluctuaciones más importantes que para la dirección principal. El resultado de esto es que el sonido proveniente de los costados no sólo estará más atenuado que el que proviene del frente (precisamente por la direccionalidad) sino que se verá más afectado en frecuencia, es decir estará "coloreado" (por ejemplo, tendrá tendencia a enfatizar los graves, u otras frecuencias específicas, produciendo respectivamente un sonido más sordo o algo metálico), como se muestra en el ejemplo de la Figura 8.8. En los mejores micrófonos, este detalle es tenido en cuenta, de modo que el patrón polar resulte más uniforme con la frecuencia, con lo cual el sonido proveniente de los costados sólo sonará más débil, y no además distorsionado.

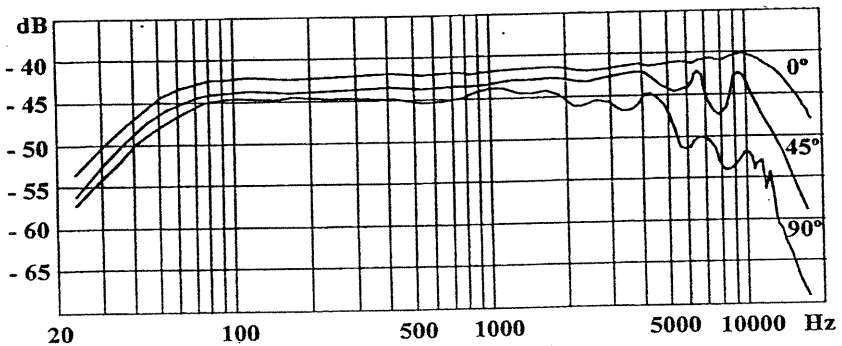


Figura 8.8. Respuesta en frecuencia de un micrófono direccional típico para diferentes ángulos respecto a la dirección principal: 0°, 45° y 90°. Según se puede apreciar, las irregularidades en alta frecuencia en la señal, mayores, introduciendo una mayor distorsión de frecuencia en la señal.

8.5. Micrófonos dinámicos

Existen varios mecanismos de conversión de energía sonora en energía eléctrica utilizados en los micrófonos. Los más habituales corresponden a los micrófonos dinámicos y los micrófonos capacitivos.

Los micrófonos dinámicos, también denominados de bobina móvil, están constituidos por una bobina con varias espiras de alambre de cobre que se desplaza en forma oscilante a lo largo de un núcleo cilíndrico de imán. La bobina es impulsada por un diafragma que vibra en concordancia con las variaciones de presión de una onda sonora (Figura 8.9). De la física se sabe que cuando una bobina se mueve dentro de un campo magnético (en este caso el del imán) se genera en sus terminales una tensión eléctrica denominada fuerza electromotriz, y éste es precisamente el principio de operación de los micrófonos dinámicos (véase el capítulo 23 para mayores detalles).

no necesito
alimentación

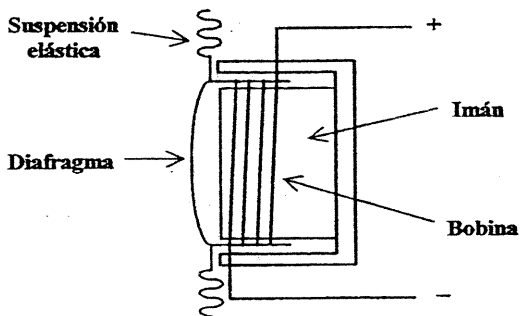


Figura 8.9. Diagrama esquemático constructivo de un micrófono dinámico o de bobina móvil. Las variaciones de presión causadas por una onda sonora imprimen movimiento al diafragma, el cual arrastra consigo a la bobina. El movimiento de la bobina respecto al imán genera una tensión eléctrica entre los terminales + y - del micrófono.

Los micrófonos dinámicos generan tensiones bastante pequeñas, del orden de 1 a 4 mV/Pa (milivolt por pascal). Para lograr mayores sensibilidades sería necesario que la bobina tuviera muchas espiras, lo cual implicaría aumentar su masa. Esto repercutiría negativamente en la respuesta en alta frecuencia, dado que a mayor masa, mayor inercia, es decir mayor dificultad para que una onda de alta frecuencia ponga en movimiento al conjunto diafragma - bobina (sería equivalente a intentar sacudir rápidamente un objeto muy pesado). De todas maneras, aún con pocas espiras (y por lo tanto baja sensibilidad) el comportamiento en alta frecuencia está limitado en general a unos 16 kHz. Actualmente, el uso de potentes imanes de neodimio permite reducir la cantidad de espiras, permitiendo en algunos modelos extender la frecuencia a la banda completa de audio.

Otra desventaja de los micrófonos dinámicos es que el denominado ruido de manipulación (es decir el ruido ocasionado al mover o tocar el micrófono) es importante, debido a dos factores: la gran inercia del conjunto diafragma - bobina y el agregado de resonancias artificiales para mejorar la respuesta en las altas y bajas frecuencias. El primer factor (inercia de la bobina) implica que al mover el micrófono la bobina tiende a permanecer inmóvil, creándose un movimiento relativo entre la bobina y el imán (ya que éste se ha desplazado junto con el cuerpo del micrófono) equivalente a que el diafragma se moviera y el imán estuviera fijo. Se genera así una tensión similar a la que produciría un ruido acústico. Esta tensión indeseada se denomina ruido eléctrico. El segundo factor (agregado de resonancias) implica que el ruido que se produce al tocar el micrófono se ve amplificado, especialmente en baja frecuencia, generando también ruido eléctrico.

La ventaja principal de este tipo de micrófonos es su robustez y tolerancia a condiciones adversas de operación, como variaciones de temperatura o humedad, grandes niveles de presión sonora, golpes y sacudidas, etc., por lo cual son especialmente aptos para el sonido en vivo. Otra ventaja es que no requieren fuentes de alimentación propias para generar señal eléctrica en respuesta a un sonido.

8.6. Micrófonos capacitivos

Los micrófonos capacitivos (también denominados de capacitor, de condensador, o electrostáticos) se basan en la utilización de un campo eléctrico en lugar de un campo magnético. Están formados por un diafragma muy delgado (típicamente, 5 micrones de espesor) bañado en oro, y una placa posterior metálica que normalmente está perforada o ranurada (Figura 8.10). Ambos forman un condensador cuya capacidad C varía con la distancia entre sí, y esta distancia varía al vibrar el diafragma impulsado por las variaciones de presión de la onda sonora. La ecuación fundamental de un condensador es

$$V = \frac{q}{C}$$

donde V es la tensión en sus terminales, q es la carga eléctrica en sus terminales, y C es el valor de la capacidad asociada al condensador. Si de alguna manera se consigue cargar las placas del condensador con una carga fija q , al variar la capacidad C variará también la tensión V que se mide entre los terminales. Este es el principio físico de funcionamiento de los micrófonos capacitivos.

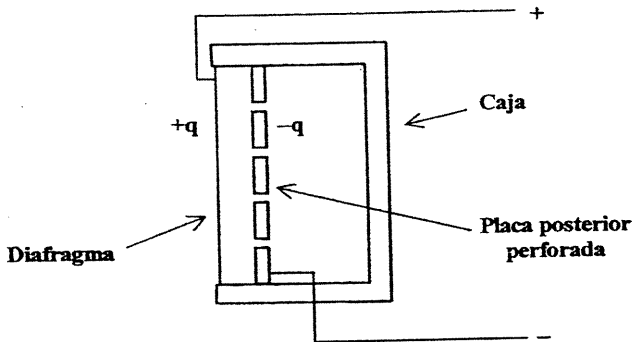


Figura 8.10. Diagrama esquemático constructivo de un micrófono de condensador. Las variaciones de presión causadas por una onda sonora imprimen movimiento al diafragma, y al variar consecuentemente la distancia entre éste y la placa posterior, varía también la capacidad del condensador formado por ambos. Si previamente se ha aplicado una carga eléctrica a ambas placas, la variación de capacidad implicará una variación de tensión eléctrica v entre los terminales $+$ y $-$ del micrófono.

La manera de cargar las placas del condensador es por medio de una polarización externa, lo cual se logra conectando el micrófono a una fuente de tensión constante a través de una resistencia, como se muestra en la Figura 8.11. Esta fuente puede ser o

bien una pila o batería incorporada al propio cuerpo del micrófono, o una fuente remota ubicada en la consola o en el preamplificador, denominada **fente fantasma (phantom power)**. Esta fuente puede tener un valor comprendido entre 1,5 V y 48 V según el modelo de micrófono.

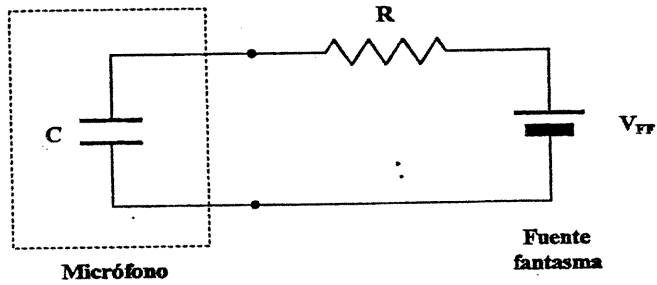


Figura 8.11. Polarización de un micrófono capacitivo mediante una fuente externa o **fente fantasma (phantom power)**. La fuente V_{FF} suministra la carga necesaria para que las variaciones de capacidad permitan obtener variaciones de tensión.

En la actualidad se encuentra muy difundido un tipo de micrófono a condensador **prepolarizado**, es decir con una polarización interna intrínseca, que en principio no requiere la utilización de una fuente fantasma. Son los denominados **electret**, y se caracterizan porque una de las placas contiene una película aislante especial en la cual se han introducido, durante la fabricación, cargas eléctricas que quedan atrapadas en la estructura interna sin posibilidad de escapar.

De todas maneras, todos los micrófonos capacitivos vienen con un amplificador interno que requiere algún tipo de alimentación. Esto se debe a que la impedancia interna (ver próximo párrafo) de un micrófono capacitivo básico es demasiado alta, por lo cual se utiliza un preamplificador sencillo (habitualmente formado por un transistor de efecto de campo) para reducir la impedancia (amplificador de corriente). Esto es necesario para minimizar la captación de ruidos en los cables largos.

8.7. Impedancia

La impedancia interna de un micrófono está vinculada con su modelo circuital o modelo eléctrico, que está constituido por una fuente de tensión y una impedancia, como se muestra en la Figura 8.12. Existen micrófonos de alta impedancia (superior a los 10.000 Ω , es decir 10 k Ω) y de baja impedancia (menor de 500 Ω). En sonido profesional se utilizan casi exclusivamente los micrófonos de baja impedancia, porque son menos ruidosos, y ofrecen menos dificultades para el cableado, en especial cuando están involucradas grandes distancias (algunas decenas de metros), como suele suceder en el sonido en vivo. El nivel de la tensión de salida es, en general, muy pequeño, (del orden de algunos μV hasta unos 100 mV), especialmente en los micrófonos de baja impedancia, razón

por la cual se requiere utilizar **preamplificadores** para elevar la tensión hasta el nivel normalmente requerido por las mezcladoras (consolas) de audio. Dichos preamplificadores por lo común están incorporados en las consolas de mezcla, y aparecen en las entradas de micrófono. *No deben confundirse estos preamplificadores con los amplificadores de conversión de impedancia incluidos en los micrófonos capacitivos (tanto los no polarizados como los prepolarizados o electret).*

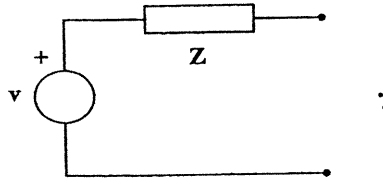


Figura 8.12. Modelo eléctrico de un micrófono, formado por una fuente de tensión v y una impedancia Z .

La impedancia de entrada de los preamplificadores debe ser mucho mayor que la del micrófono, para no ocasionar un efecto de divisor de tensión, lo cual redundaría en una disminución de la tensión efectiva en la entrada del preamplificador. El efecto de divisor de tensión se da siempre en conexiones como la indicada en la Figura 8.13. Resulta

$$V_{amp} = \frac{Z_{entrada}}{Z + Z_{entrada}} v,$$

y por lo tanto cuanto más grande sea $Z_{entrada}$ menor será la disminución de la tensión que recibe el preamplificador. En la práctica la impedancia de los micrófonos suele ser de unos 200Ω y la de las entradas de micrófono de alrededor de 1000Ω .

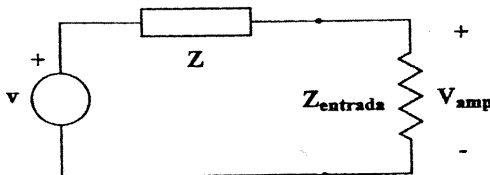


Figura 8.13. Modelo eléctrico de la conexión entre un micrófono y su preamplificador. El micrófono tiene una impedancia Z y el preamplificador tiene una impedancia de entrada $Z_{entrada}$.

8.8. Ruido

En los micrófonos hay dos mecanismos de producción de ruido. El más evidente es la captación del ruido ambiental, y obedece al mismo principio de conversión de energía sonora en energía eléctrica que tiene lugar para los sonidos útiles. La reducción de este ruido está ligada a la reducción del propio ruido ambiente, y al aprovechamiento del patrón direccional para reducir los ruidos que provienen de direcciones distintas de la de la señal útil (una voz o instrumento).

El otro mecanismo es el característico de cualquier componente de un circuito, es decir el ruido eléctrico. Es un ruido intrínseco del micrófono, y aparecería aún ubicándolo en una cámara totalmente insonorizada (silenciosa). Este ruido sólo puede reducirse (pero no eliminarse) diseñando el micrófono de modo de que posea muy baja impedancia (por ejemplo 100Ω), y además utilizando en su fabricación materiales de gran calidad y procesos de manufactura sumamente refinados. Es interesante saber que un micrófono de 100Ω tiene como mínimo un ruido eléctrico de $0,18 \mu\text{V}$, y que este ruido se duplica cada vez que la impedancia se cuadruplica. En la especificación de un micrófono sólo tiene sentido especificar el ruido eléctrico, ya que el otro depende del nivel de ruido acústico del ambiente en donde se usa el micrófono, y por lo tanto no es atribuible del micrófono.

Existen dos formas de especificar el ruido eléctrico. La primera consiste en asociarlo a un nivel de presión sonora equivalente, por ejemplo 17 dB . En otras palabras, se asimila el ruido eléctrico, que corresponde a un valor pequeño de tensión, a la respuesta hipotética de un micrófono sin ruido propio ante un ruido ambiente del valor indicado. Esta especificación sugiere cuál es el mínimo nivel sonoro para el cual tendrá sentido utilizar este micrófono. Si se intentara captar con ese micrófono un sonido de menor nivel, el ruido propio del micrófono sería más intenso que el sonido a captar, con un serio deterioro de la calidad sonora, o hasta el franco enmascaramiento de la señal por el ruido. *(Obsérvese que el solo hecho de que un sonido sea menor que otro no necesariamente implica que sea enmascarado por aquél, aunque sí severamente degradado. Como regla general, un sonido debe estar en la misma banda de frecuencias y entre 15 y 20 dB por debajo de otro para resultar enmascarado.)*

La segunda forma de especificar el ruido de un micrófono, es a través del concepto de relación señal/ruido. Se define la relación señal/ruido como el cociente entre la señal y el ruido, y a menudo se la expresa en dB :

$$S/R|_{\text{dB}} = 20 \log_{10} \frac{\text{señal}}{\text{ruido}}$$

Si, por ejemplo, la señal es de 2 mV y el ruido de $1 \mu\text{V}$ ($= 0,001 \text{ mV}$), resulta

$$S/R|_{\text{dB}} = 20 \log_{10} \frac{2 \text{ mV}}{0,001 \text{ mV}} = 20 \log_{10} 2000 = 66 \text{ dB}$$

El mismo valor se obtendría si, por ejemplo, el nivel de la señal fuera de 200 mV y el nivel de ruido de $100 \mu\text{V}$ ($= 0,1 \text{ mV}$).

Como puede deducirse, para que esta especificación tenga sentido es muy importante incluir el nivel de la señal que se está utilizando en la especificación. Así, la misma

cifra de 66 dB representaría una muy buena relación señal/ruido si la señal utilizada fuera muy pequeña, y en cambio sería muy mala si la señal fuera la máxima que puede manejar el micrófono.

Normalmente la especificación de la señal no se da en mV, sino en Pa (presión) o en dB (nivel de presión sonora), y se incluye la frecuencia de la señal. Una posible especificación podría ser:

$$S/R: 50 \text{ dB a } 1 \text{ kHz; } 0,1 \text{ Pa}$$

o bien

$$S/R: 50 \text{ dB a } 1 \text{ kHz; } 74 \text{ dB NPS}$$

(Obsérvese que estas especificaciones son idénticas, ya que una presión de 0,1 Pa corresponde a un nivel de presión sonora de 74 dB). En el segundo caso es fácil obtener el nivel de presión sonora equivalente al ruido eléctrico:

$$NPS_{\text{equivalente}} = 74 \text{ dB} - 50 \text{ dB} = 24 \text{ dB}$$

Esto significa que el límite inferior de utilidad del micrófono es de 24 dB de nivel de presión sonora.

8.9. Distorsión

Otra especificación de interés en los micrófonos es la distorsión. La distorsión se diferencia del ruido en que es una deformación de la onda, mientras que el ruido es una señal independiente que se agrega a la señal. Cuando la señal es una onda senoidal, la distorsión se manifiesta como la aparición de cierta cantidad de armónicos. Así, si la señal (en este caso una presión sonora) era de 500 Hz, por ejemplo, la señal resultante (en este caso la tensión eléctrica generada por el micrófono) contendrá principalmente una onda senoidal de 500 Hz pero también contendrá armónicos de dicha frecuencia, vale decir, senoides de frecuencia 1000 Hz, 1500 Hz, 2000 Hz, etc. Se define la **distorsión total armónica (THD)** como el cociente entre el valor eficaz de los armónicos generados por la distorsión y el valor eficaz de la fundamental, y se suele expresar en porcentaje. Por ejemplo, supongamos que se expone un micrófono a un sonido senoidal, y que como resultado produce una tensión que tiene 5 mV de primera armónica y 0,05 mV de las restantes armónicas, entonces la distorsión total armónica será:

$$THD = \frac{0,05 \text{ mV}}{5 \text{ mV}} = 0,01 = 1\%$$

Dado que el fenómeno de la distorsión se da normalmente para niveles altos de señal, la especificación se suele dar asociada con el máximo nivel de presión sonora que admite el micrófono. Una posible especificación podría ser:

$$THD: 1\% \text{ a } 125 \text{ dB NPS}$$

Otra forma de expresar lo mismo sería:

Máximo NPS: 125 dB a 1% THD .

El valor máximo, salvo que se indique lo contrario, representa un nivel operativo, y no un nivel de seguridad. La diferencia entre ambos es que el nivel operativo es un nivel para el cual el micrófono todavía está funcionando razonablemente bien, mientras que el nivel de seguridad es aquel superado el cual el micrófono puede deteriorarse. No superar este valor asegura que el dispositivo no se destruye, pero no que funcione satisfactoriamente. Si se reduce la señal hasta el nivel operativo, el funcionamiento volverá a ser correcto.

8.10. Conexión balanceada

A las especificaciones relativas a la performance del micrófono, se agregan otras de carácter más administrativo que operativo. Una de ellas es el tipo de conexión y cableado. Existen dos tipos de conexión clásicas. La más básica es la conexión simple o no balanceada, que consiste en dos conductores que unen la cápsula del micrófono al amplificador. Por regla general uno de los conductores rodea al otro formando un mallado o blindaje electrostático cuya finalidad es minimizar la captación de ruido eléctrico por efecto antena. Dicho conductor se conecta normalmente a la masa de la alimentación del amplificador. La otra conexión es la de tipo balanceado. En este caso se utilizan tres conductores. Uno de ellos forma también un mallado y actúa como blindaje para los otros, que van por dentro. La característica más importante es que estos conductores llevan señales de signo opuesto con respecto a la malla. Es decir que si en determinado momento uno tiene una tensión de 1 mV, el otro tendrá una tensión de -1 mV. La finalidad de esto es reducir mucho más la captación de ruido eléctrico por los cables, particularmente el ruido inducido por acoplamiento electromagnético, que no es fácil de controlar debido a las dificultades para implementar un adecuado blindaje magnético. Esta inmunidad se logra utilizando como preamplificador un amplificador diferencial, es decir un amplificador que amplifica la diferencia entre las tensiones de sus dos entradas (Figura 8.14).

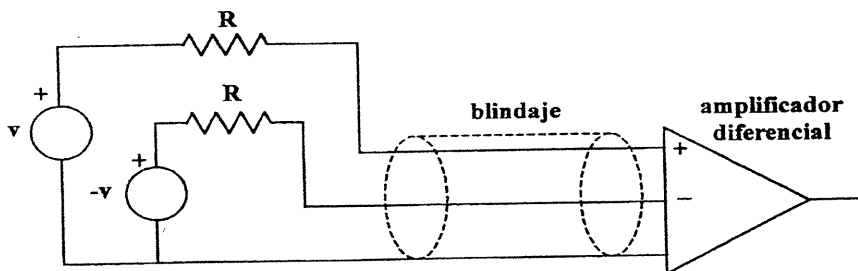


Figura 8.14. Estructura de una conexión balanceada. El cable de conexión tiene tres conductores, uno de los cuales es el blindaje.

Supongamos que $v = 1 \text{ mV}$, y que a causa del efecto antena se genera en cada cable de conexión una tensión de ruido de 10 mV. Entonces la tensión que llega al ter-

minal no inversor (+) es de 11 mV, mientras que la que llega al terminal inversor (-) es de 9 mV. El amplificador diferencial amplifica la diferencia,

$$11 \text{ mV} - 9 \text{ mV} = 2 \text{ mV} ,$$

que es *la misma tensión* que se obtendría si los cables no captaran ruido:

$$1 \text{ mV} - (-1 \text{ mV}) = 2 \text{ mV} .$$

Debe observarse que el hecho de que se genere en los dos terminales *la misma* tensión (en el ejemplo, 10 mV) se debe a que los dos conductores se encuentran físicamente muy próximos entre sí, y por lo tanto están prácticamente a la misma distancia del elemento que ocasiona el campo electrostático (un cable de distribución domiciliar de 220 V, un tubo fluorescente, un motor, etc.). Esta tensión idéntica en ambos conductores balanceados se denomina **tensión de modo común**.

NOTA: Podría preguntarse por qué no sucede lo mismo en el cable de masa (la malla). Ello se debe a que el cable de masa está siempre conectado a circuitos de *muy baja impedancia*, lo cual atenúa mucho la captación de ruido.

El conector para lograr la conexión balanceada se denomina habitualmente XLR, aunque también se lo conoce como Cannon, que fue la primera marca comercial que lo introdujo.

8.11. Fuente fantasma

Finalmente, en los micrófonos capacitivos (a condensador), se especifica la alimentación, que puede ser con pila o batería ubicada en el mismo cuerpo del micrófono, o

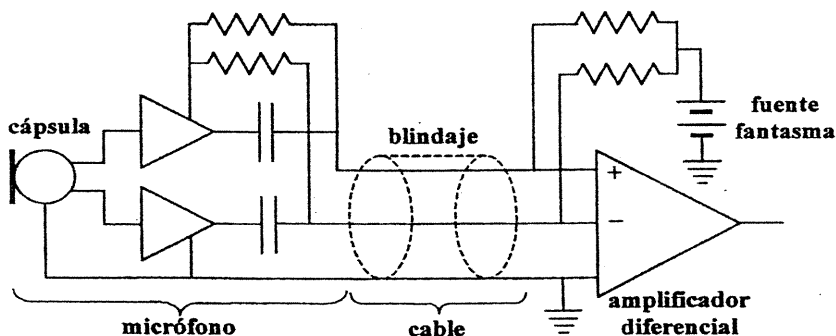


Figura 8.15. Esquema del conexionado de la fuente fantasma. La salida de la cápsula (condensador) se conecta a unos amplificadores cuya finalidad es la de reducir la impedancia naturalmente muy alta del condensador. El usuario sólo debe preocuparse por el cableado, ya que la conexión interna del micrófono y del amplificador diferencial de entrada de la consola ya vienen ensambladas de fábrica.

con fuente fantasma o remota. La fuente fantasma se utiliza en los micrófonos con conexión balanceada, y su estructura circuital se muestra en la Figura 8.15. Las salidas de la cápsula (condensador) ingresan a unos amplificadores cuya única finalidad es la de reducir la impedancia excesivamente alta del condensador hasta los niveles requeridos (según vimos, algunos cientos de ohm). Estos amplificadores, así como la propia cápsula, requieren alimentación, la cual llega de la consola por medio de sendas resistencias (como la potencia requerida es pequeña, estas resistencias no implican un inconveniente). Luego siguen dos capacitores cuya finalidad es eliminar las corrientes continuas, y dejar pasar sólo las frecuencias de audio. Estos capacitores son de valor mucho más alto que los de la cápsula, por lo que su impedancia es muchísimo menor, y por lo tanto no destruyen el efecto de los amplificadores mencionados. Los capacitores se conectan (por medio del conector XLR) al cable y luego a la consola, donde las dos señales complementarias ingresan al amplificador diferencial. La fuente fantasma, que hoy en día se encuentra habitualmente incorporada a la consola, aplica la misma tensión a las líneas de señal por medio de otro par de resistencias. Dado que el amplificador es diferencial, la tensión agregada no influye por ser igual en ambas entradas (por la misma razón por la cual la conexión balanceada rechaza el ruido de modo común).

Debido que no todos los micrófonos son de condensador, es lógico preguntarse qué sucede cuando se conecta un micrófono dinámico, por ejemplo, a la entrada de micrófono con fuente fantasma. Si el micrófono es balanceado (se supone que todos los de buena calidad profesional lo son), no sucede nada, ya que se aplica la misma tensión a ambas partes de la cápsula, con lo cual la tensión aplicada es nula. Pero si el micrófono fuera no balanceado, la tensión estaría aplicada directamente a la bobina, lo cual sería potencialmente muy perjudicial para el micrófono, pudiendo fácilmente destruirse. Existen algunos micrófonos dinámicos pseudobalanceados, es decir que por compatibilidad de conexión tienen conector XLR, pero no son realmente balanceados. Habitualmente, las consolas tienen un interruptor que permite conectar o desconectar la fuente fantasma, por lo cual, ante la duda, siempre convendrá desconectarla si no se están empleando micrófonos de condensador. Lamentablemente, no todas las consolas tienen una fuente fantasma por cada canal, por lo cual este interruptor habilita o deshabilita la fuente fantasma para todos los canales simultáneamente.

Antes de abandonar el tema de los micrófonos, comentemos que existen otros tipos de micrófonos, como los piezoeléctricos, los de cinta y los de carbón. Los de cinta han encontrado algún lugar en el audio profesional, pero en general los más utilizados son los ya estudiados de condensador y dinámicos.

Capítulo 15

Audio Digital

15.1. Introducción

Las técnicas digitales han cobrado en las últimas décadas una importancia fundamental en el desarrollo de nuevas tecnologías para la generación, el procesamiento, el almacenamiento y el análisis del sonido. Ello ha sido posible gracias al avance vertiginoso de la microelectrónica y su aplicación a la producción de dispositivos poderosos y complejos capaces de manejar y transformar cada vez más con mayor precisión y rapidez la enorme cantidad de información contenida en el sonido.

Una de las primeras consecuencias de la aplicación de la tecnología digital al audio fue el desarrollo de sistemas de almacenamiento del sonido de gran confiabilidad, inalterabilidad y fidelidad. Otra fue el gran impulso al desarrollo de instrumentos musicales electrónicos de gran complejidad y versatilidad. La tercera consecuencia fue el desarrollo y aplicación de técnicas para el procesamiento de la señal sonora, que permitieron no sólo el mejoramiento de procesos que antes se llevaban a cabo analógicamente sino también la introducción de nuevos procesos, entre los cuales se encuentran una gran cantidad de efectos tales como retardos, modulaciones, reverberaciones y espacializaciones de gran realismo y naturalidad, cuya implementación analógica sería mucho más costosa y por lo tanto destinada a un mercado mucho más restringido.

La idea básica detrás del audio digital es la de *representar el sonido por medio de números* ("digital" viene de *dígito*, es decir número). Aún antes de profundizar el análisis, es fácil ver que esto tiene varias ventajas. En primer lugar, *se elimina el problema de la alterabilidad de la información. Es mucho más fácil guardar un número que la magnitud física que ese número representa.* Por ejemplo, si quisiéramos guardar una varilla de 57,235 cm de longitud, tendríamos serias dificultades. Efectivamente, la dilatación a causa de la temperatura, o cualquier partícula de polvo que se adhiriera a sus extremos, o simplemente el desgaste, podrían ocasionar un error. Esto, que es válido para la longitud de una varilla, lo es más para el campo magnético almacenado en una cinta grabada.

En segundo lugar, *existen algoritmos (métodos de cálculo) para realizar digitalmente no sólo todos los tipos de procesamiento utilizados en el audio tradicional, como la amplificación, la mezcla, la modulación, el filtrado, la compresión y expansión, etc., sino muchos otros más, entre los cuales se encuentran los retardos, los sincronismos, los desplazamientos de frecuencia, la generación de sonidos por diversos procedimientos, etc.* Estos algoritmos pueden implementarse en una computadora de propósito general o bien en dispositivos específicos llamados **procesadores digitales de señal (DSP)**.

En tercer lugar, el reemplazo de los procesadores analógicos por sus equivalentes digitales permite evitar la degradación de la señal a causa del ruido analógico, lo cual es conveniente dado que el ruido analógico es muy difícil de eliminar.

15.2. Numeración binaria

Dado que todos los sistemas digitales se basan en la numeración binaria, antes de comenzar a describir los procesos básicos de muestreo y digitalización del sonido nos referiremos brevemente a esa numeración. En la numeración decimal (el sistema que empleamos habitualmente), se utilizan diez símbolos (los dígitos 0, 1, 2, ..., 9) en un sistema *posicional* para representar las sucesivas cantidades. Esto significa que cada nueva cifra que se agrega tiene un peso 10 veces mayor que la que se encuentra a su derecha. Por ejemplo,

$$27 = 2 \times 10 + 7,$$

$$306 = 3 \times 10^2 + 0 \times 10 + 6.$$

En la numeración binaria, se utilizan sólo dos símbolos (los dígitos 0 y 1), también en un sistema *posicional*, sólo que ahora cada nueva cifra tiene un peso sólo 2 veces mayor que la anterior. Por ejemplo,

$$101 = 1 \times 2^2 + 0 \times 2 + 1 = 5,$$

$$11011 = 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2 + 1 = 27.$$

En esta numeración, el 1 tiene características similares al 9 del sistema decimal, es decir, una vez que llegamos al 1 debemos agregar una nueva cifra 1 y cambiar la primera por 0. En la Tabla 15.1 se muestra la conversión de decimal a binario para los números del 0 al 15.

La razón por la que se utilizan los números binarios es porque eléctricamente es muy fácil codificar los 0's y los 1's. Basta utilizar un nivel de tensión alto (5 V) para un 1 y un nivel de tensión bajo (0 V) para un 0. Esto hace que la representación sea extremadamente insensible al ruido. En efecto, la señal seguiría siendo recuperable aún en presencia de un ruido de 2 V, que corresponde a una relación señal/ruido tan baja como $20 \log 5/2 = 8 \text{ dB}$ (inadmisible si el sistema fuera analógico).

15.3. Muestreo

Pasemos ahora al concepto de muestreo (*sampling*). Las señales acústicas (y por lo tanto las señales eléctricas que las representan) varían en forma continua, lo que significa que en un intervalo de tiempo dado, por pequeño que sea, existen *infinitos* valores diferentes. Sin embargo, a los efectos del mensaje auditivo, no hace falta tanta información. Primero, porque el oído no tiene tanta *discriminación en el tiempo*, y segundo porque tampoco tiene tanta *discriminación en la amplitud* como para distinguir valores que por estar muy próximos en el tiempo difieren muy poco en amplitud. No solamente no

Tabla 15.1. Conversión del sistema binario al decimal.
Los ceros de más a la izquierda son opcionales.

Decimal	Binario
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111

hace falta tanta información sino que desde el punto de vista práctico tampoco es conveniente ni posible manejarla. Entonces surge el concepto de muestreo (sampling). Muestrear una señal significa *reemplazar la señal original por una serie de muestras tomadas a intervalos regulares*. La frecuencia con la que se toman las muestras se denomina frecuencia de muestreo, f_M , y el tiempo entre muestras, periodo de muestreo, T_M . Se cumple que

$$f_M = \frac{1}{T_M}$$

En la Figura 15.1 se ilustra el proceso de muestreo. En la grafica superior se grafica la onda original y los instantes de muestreo, y en la figura de abajo se indican las muestras.

15.4. Frecuencia de muestreo

Es intuitivamente evidente que la frecuencia de muestreo debe ser bastante alta, ya que entonces se logra un grado de detalle mucho mayor, lo cual significa que el sonido será reproducido con mayor fidelidad al original. En realidad existe un criterio que debe cumplirse obligatoriamente en todo proceso de muestreo, y es que *la frecuencia de muestreo debe ser mayor que el doble de la máxima frecuencia presente en la señal*. Es decir

$$f_M > 2f_{\text{máx}}$$

Esto es consecuencia de un teorema llamado **Teorema del muestreo**, que dice que una señal muestreada puede recuperarse totalmente sólo si fue muestreada cumpliendo con el criterio anterior. La frecuencia $f_M/2$ se denomina **frecuencia de Nyquist**.

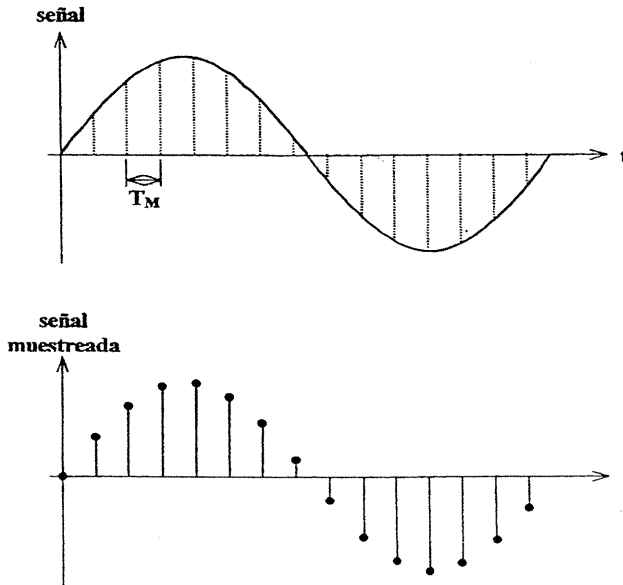


Figura 15.1. Efecto del proceso de muestreo sobre una onda senoidal. La frecuencia de muestreo es en este caso 14,7 veces mayor que la frecuencia de la onda.

Es importante comprender que la frecuencia máxima que aparece en la fórmula anterior no se refiere sólo a la máxima frecuencia *de interés*, sino en realidad a la máxima frecuencia que *efectivamente* aparece en la señal a muestrear, aunque dicha frecuencia provenga de un ruido de alta frecuencia que contamina la señal. En caso de no cumplirse el criterio, al intentar recuperar la señal aparecerán componentes de frecuencia en la banda útil. Para verlo, supongamos que muestreamos con una frecuencia de 40 kHz una señal de audio, y que aparece un ruido (inaudible) de 35 kHz superpuesto a la señal, situación ilustrada en la **Figura 15.2**. Como consecuencia del proceso de muestreo y posterior reconstrucción de la señal, aparece una frecuencia de 5 kHz que *no se encontraba presente en la señal original*. Esta frecuencia, que sustituye a la original de 35 kHz, se denomina *alias* de aquella. Obsérvese especialmente que la frecuencia original (35 kHz) no producía sensación audible, pero la nueva frecuencia, no sólo es audible sino que está cerca de la región de máxima sensibilidad del oído y por lo tanto se percibirá como un silbido notorio y molesto.

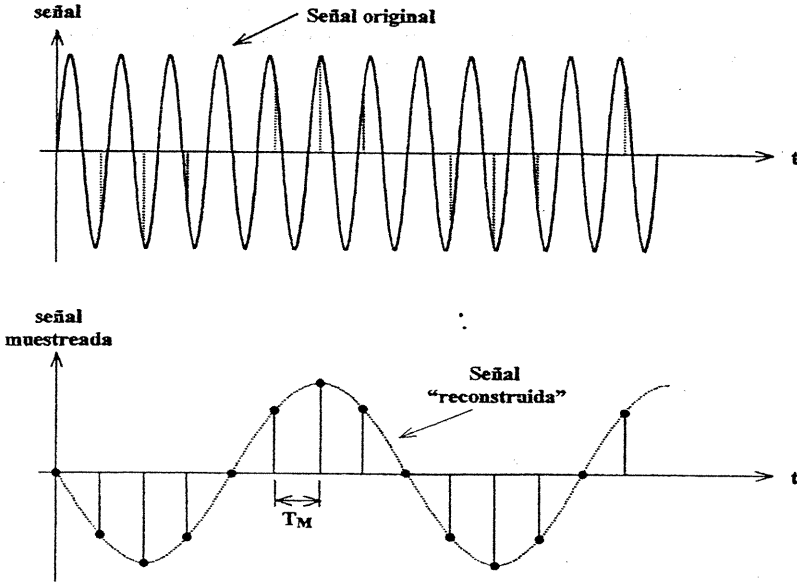


Figura 15.2. Efecto del muestreo con una frecuencia menor que el doble de la máxima frecuencia contenida en la señal. Una señal de 35 kHz se muestrea con una frecuencia de 40 kHz; al intentar reconstruirla, aparece una frecuencia *alias* de 5 kHz.

El ejemplo anterior nos está señalando que si pretendemos que la señal pueda reconstruirse correctamente después del proceso de muestreo es imprescindible eliminar toda frecuencia espuria que caiga más allá del espectro de audio, es decir por encima de 20 kHz. Se utiliza para ello un filtro pasabajos de pendiente muy abrupta en la banda de corte (96 dB/octava ó más), denominado filtro antialias (en inglés antialiasing filter).

La elección como frecuencia de muestreo estándar de 44,1 kHz para audio digital obedece precisamente a este problema de las frecuencias alias y la consecuente necesidad de un filtro antialias. Si imponemos una frecuencia máxima de 20 kHz para el audio de alta calidad, el filtro antialias deberá tener su frecuencia de corte en 20 kHz, y como su caída es rápida pero no *infinitamente* rápida, recién después de los 22 kHz se puede considerar que las señales espurias han quedado reducidas a niveles despreciables (Figura 15.3). Por ello se ha adoptado una frecuencia de algo más del doble, es decir 44,1 kHz (el valor *exacto* de 44,1 kHz en lugar de 44 kHz surgió en los comienzos de la grabación digital en cinta de video, para compatibilizar la norma de audio con la de video).

Un inconveniente de los filtros antialias es su gran complejidad y el hecho de que no son del todo inofensivos para la señal dentro de la banda de paso (en este caso la de audio). Aunque el filtro afecte sólo imperceptiblemente la amplitud de la señal en dicha banda, afecta de un modo apreciable la fase, lo cual puede alterar la imagen estéreo.

Además, afecta las señales rápidamente variables agregando pequeños transitorios de frecuencias próximas a la de corte. Por esa razón, otros formatos de audio digital utilizan frecuencias de muestreo más altas, como 48 kHz (el DAT, por ejemplo), que requieren filtros menos complejos. En realidad la frecuencia de 44,1 kHz obedece limitaciones tecnológicas propias de la época en que surgió el compact disc.

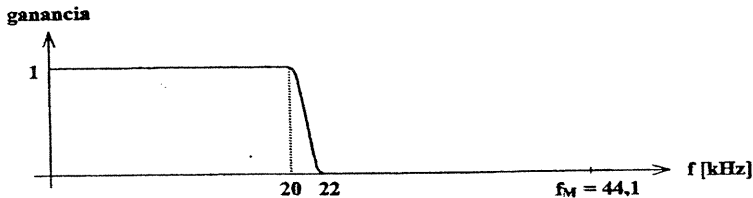


Figura 15.3. Respuesta en frecuencia de un filtro antialias utilizado para audio digital de alta calidad.

15.5. Digitalización

Una vez tomada cada muestra es necesario almacenarla, y para ello debe ser transformada en un número, más específicamente, en un número binario. Esta función la cumple un dispositivo denominado **convertor analógico-digital (A/D)**, que convierte valores de tensión en números binarios.

Consideremos el ejemplo de la Figura 15.4, en el cual utilizamos números binarios de 3 dígitos. Dado que un dígito binario se denomina **bit** (del inglés, **binary digit**), estaremos utilizando, por lo tanto, números de 3 bits. Es fácil ver que existen 8 ($= 2^3$) números de 3 bits: 000, 001, 010, 011, 100, 101, 110 y 111. Para representar los diversos valores de tensión que pueden tomar las muestras, dividimos el rango de variación de la señal en 8 niveles, y aproximamos cada muestra al nivel inmediato inferior.

En la parte central de la Figura 15.4 se comparan las muestras exactas (puntos vacíos), y las muestras digitalizadas (puntos llenos), aproximándolas de la manera comentada. Vemos que el error máximo que se comete es de una división, que a su vez corresponde a 1 bit. La forma de onda reconstruida difiere considerablemente de la original debido a que una resolución de 3 bits es muy pequeña.

En el ejemplo anterior adoptamos, en forma arbitraria, una resolución de 3 bits. El resultado fue, como pudo observarse, bastante deficitario, ya que la onda reconstruida estaba muy distorsionada. Sería interesante disponer de algún criterio más sistemático para seleccionar la resolución requerida.

El problema es similar al de decidir cuántos dígitos decimales se necesitan para representar una longitud dada con suficiente precisión. Por ejemplo, si quisiéramos representar la longitud de objetos menores de 1 m con una precisión milimétrica, necesitaríamos 3 dígitos decimales, ya que dichos objetos podrían medir entre 0 y 999 mm. Si en cambio quisiéramos una precisión de décimas de milímetro, necesitaríamos 4 dígitos, ya que los objetos podrían medir entre 0 y 9.999 décimas de mm.

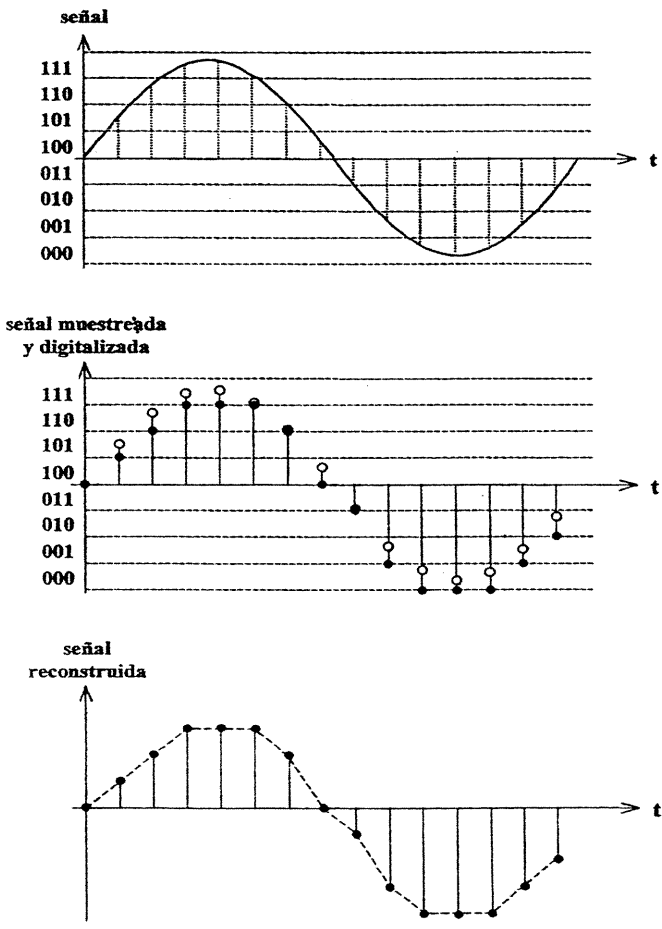


Figura 15.4. Efecto del proceso de muestreo y digitalización sobre una onda senoidal. La resolución es de 3 bits y la frecuencia de muestreo 14,7 veces mayor que la frecuencia de la onda. En la figura central los puntos vacíos representan las muestras exactas y los puntos llenos las muestras digitalizadas. Abajo se muestra la señal reconstruida.

En audio, el criterio para determinar la "precisión" es la relación señal/ruido. Analicemos desde este punto de vista el ejemplo de la Figura 15.4. Dejando de lado el ruido propio que pudiera contener la señal, un efecto colateral de la digitalización es la aparición de un error, el cual puede asimilarse a un ruido. Este ruido se conoce como ruido de digitalización. Bajo esta interpretación, el máximo valor pico a pico de la señal es proporcional a 8, y el máximo valor pico a pico de ruido es proporcional a 1. Entonces, la relación señal a ruido es $8/1 = 8$, que expresada en dB es

$$S/R = 20 \log_{10} \frac{8}{1} = 18 \text{ dB} .$$

Si tenemos en cuenta que en audio de alta fidelidad se manejan hoy en día relaciones señal/ruido mayores de 90 dB, podemos comprender por qué una resolución de 3 bits es por completo insuficiente.

Supongamos ahora que aumentamos la resolución a 4 bits. Dado que ahora hay 16 posibles valores en lugar de 8, la relación señal a ruido en dB será ahora

$$S/R = 20 \log_{10} \frac{16}{1} = 24 \text{ dB} .$$

Vemos que se ha incrementado en 6 dB. Esto puede interpretarse así: si bien la amplitud de la señal no cambió, al duplicar la cantidad de niveles, cada nivel se redujo a la mitad, por lo cual el ruido de digitalización también se redujo a la mitad. Entonces la relación señal a ruido se duplica, y una duplicación equivale a un incremento de 6 dB. Si ahora incrementáramos la resolución nuevamente en 1 bit, llevándola a 5 bits, observaríamos que nuevamente el ruido se reduciría a la mitad, por lo que la relación señal/ruido experimentaría otro incremento de 6 dB.

Podemos obtener una expresión general para la relación señal a ruido. Si adoptamos una resolución de n bits, donde n es cualquier número entero, resulta

$$S/R_{n \text{ bits}} = 6 \cdot n \text{ dB} .$$

Aplicando esta fórmula a la resolución estándar de 16 bits utilizada en los formatos más populares de almacenamiento de sonido digital, resulta una relación señal/ruido de 96 dB. Esta relación señal/ruido es, en condiciones normales, suficiente para crear contrastes dinámicos imponentes. En efecto, tengamos en cuenta que muy rara vez se tiene en la música un nivel sonoro de más de 110 dB (el cual es realmente ensordecedor y para nada recomendable). Si restamos a este valor 96 dB, obtenemos 14 dB, nivel sonoro que probablemente pocas personas tengan el privilegio de haber "escuchado", ya que aún en condiciones de gran silencio durante la noche, en una habitación interior, normalmente es difícil bajar de los 20 dB de nivel de presión sonora.

Es necesario advertir que aunque un sistema funcione con formato de audio digital de 16 bits, su relación señal/ruido *no necesariamente será de 96 dB*. Esto se debe a que en los diversos componentes analógicos que forman parte de todo dispositivo se genera ruido que se agrega al ruido de digitalización. En los equipos de bajo costo, como por ejemplo los discman o los minicomponentes, la electrónica de baja costo (y baja calidad) utilizada en su fabricación es particularmente ruidosa y la relación señal/ruido real es bastante menor que 96 dB.

15.6. Reconstrucción de la señal

Veamos ahora con más detalle el proceso de reconstrucción de la señal. El procedimiento más simple consiste en obtener, mediante un conversor digital-analógico (D/A), un valor proporcional al número binario de cada muestra y mantenerlo constante hasta

que llegue una nueva muestra, es decir durante un periodo de muestreo T_M . Este procedimiento se denomina **retención simple**. En la Figura 15.5 se muestra este procedimiento para la señal digitalizada de la Figura 15.4.

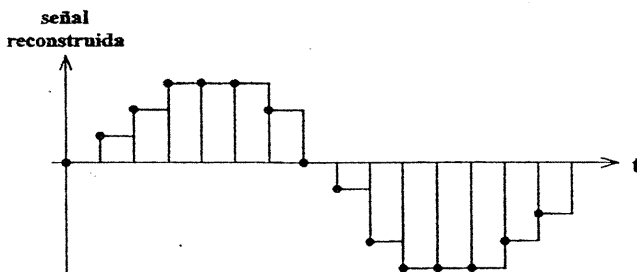


Figura 15.5. Reconstrucción de la señal digitalizada de la Figura 15.4 mediante una retención simple.

Una vez reconstruida la señal, debemos utilizar un **filtro de suavizado**, es decir un **filtro pasabajos** que quite la forma escalonada que resulta de la retención simple. Dicho filtro debe tener características similares a las del filtro antialias que habíamos introducido para la digitalización, es decir, debe ser bastante abrupto para eliminar casi por completo las frecuencias superiores a los 20 kHz, y en cambio permitir pasar las inferiores. Nuevamente, un filtro de ese tipo por un lado es complejo y por el otro produce distorsiones de fase. Para resolver esta situación se ha introducido el concepto de **sobremuestreo (oversampling)**.

El **sobremuestreo** consiste en intercalar, entre las muestras de la señal *realmente* obtenidas o almacenadas, otras "muestras" calculadas por interpolación. Así, por ejemplo, un sobremuestreo por 8 agrega 7 muestras calculadas por cada muestra real. El resultado equivale a una frecuencia de muestreo 8 veces superior a la original. Si $f_M = 44,1$ kHz, entonces la nueva frecuencia de muestreo es de 352,8 kHz, la cual puede eliminarse con filtros pasabajos muchísimo más sencillos y con menos efectos sobre la fase y sobre los transitorios de la señal. El sobremuestreo se utiliza hoy corrientemente en los reproductores de compact disc, lo cual es posible porque la velocidad de la electrónica es mucho mayor que lo que era cuando recién había surgido esta nueva tecnología.

15.7. Memorias electrónicas

Introduzcamos ahora algunos conceptos relativos a las memorias electrónicas. Existen básicamente dos tipos de memorias: las memorias **RAM (random access memory)**, y las memorias **ROM (read only memory)**. Las memorias **RAM** son memorias de *lectura y escritura*, es decir en las cuales es posible escribir (guardar) números binarios, y posteriormente leerlos. Son memorias *volátiles*, es decir que una vez interrumpido el suministro de energía eléctrica (al desconectar el equipo) la información se pierde. Las memorias **ROM** son memorias de *lectura solamente*, en las cuales no es posible escribir.

En ellas hay datos fijos de fábrica, que *no se pierden al desenergizar el equipo*. Se utilizan mucho para guardar formas de onda muestreadas, como en la mayoría de los sintetizadores actuales.

En ambos tipos de memoria los datos se almacenan en posiciones sucesivas, cada una de ellas identificada con un número binario denominado *dirección*. En la Figura 15.6 se ilustra con un ejemplo la estructura de una memoria con datos de 4 bits y direcciones de 3 bits. Así, en la dirección 0, es decir 000, se encuentra almacenado el número binario 0110 (en decimal, 6), y en la dirección 3, es decir 011, se encuentra almacenado

000	001	010	011	100	101	110	111
0110	0001	1110	1101	0010	1000	0001	0010

Figura 15.6. Estructura de almacenamiento de una memoria digital. La memoria se organiza en posiciones identificadas con direcciones.

el número binario 1101 (en decimal, 13). En la Figura 15.7 se muestran las entradas y salidas principales de una memoria RAM. Cuando nos referimos a una entrada de datos de 4 bits, por ejemplo, en realidad está formada por 4 líneas o cables, y por eso tanto la entrada como la salida y la dirección se han dibujado como flechas gruesas. Las entradas

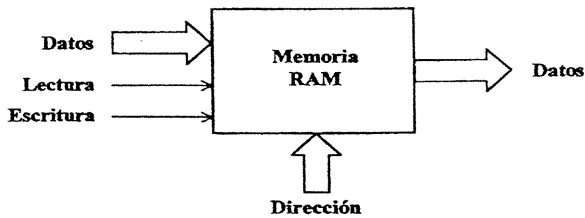


Figura 15.7. Entradas y salidas de una memoria RAM. Las flechas delgadas representan señales binarias de 1 bit de control, y las flechas gruesas son datos digitales de varios bits.

de lectura y escritura son señales binarias de 1 bit, cuya función es de control. Cuando la *entrada de escritura* toma el valor 1, el dato digital X presente en ese instante en la entrada de datos se escribirá en la dirección D indicada por la entrada de direcciones. Para leer dicho valor, bastará volver a aplicar el valor D en la entrada de direcciones y dar valor 1 a la *entrada de lectura*. El valor guardado aparecerá en la salida de datos.

Desde el punto de vista del conexionado, la única diferencia entre una memoria RAM y una ROM es que esta última no tiene entrada de escritura. En el caso de la RAM, después de escribir un dato en una dirección éste quedará almacenado allí hasta que se escriba un nuevo dato en el mismo lugar o hasta que se interrumpa el suministro de energía eléctrica al circuito.

Aunque por razones conceptuales nos hemos referido a la entrada y la salida de datos como si fueran entidades físicamente diferentes, en realidad son el mismo conjunto de líneas cuya función en cada instante depende de cuál de las entradas de control valga 1. Si ambas entradas son 0, dichas líneas se desconectan de modo de no cargar al resto del circuito.

15.8. Dither

Quando se digitalizan señales de muy bajo nivel (cercano a la resolución del conversor) el ruido de digitalización se convierte en una distorsión, cuyo efecto es más perjudicial que el de un ruido aleatorio. Por ejemplo, si se digitaliza una señal senoidal de 100 Hz y amplitud apenas menor que un escalón (Figura 15.8), se obtiene una señal

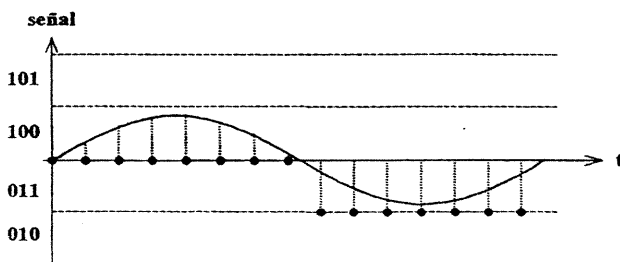


Figura 15.8. Distorsión creada al muestrear señales de bajo nivel.

que al reconstruirse será muy similar a una onda cuadrada, y por lo tanto contendrá armónicos de 300 Hz, 500 Hz, 700 Hz, etc. Si en lugar de una onda senoidal se aplicaran dos o más, aparecería, similarmente, una distorsión por intermodulación sumamente indeseable.

Una manera de evitar estos inconvenientes es aplicar una pequeña cantidad de ruido aleatorio antes del muestreo y posterior digitalización. Este ruido, cuyo valor eficaz es, normalmente, menor que un escalón, se denomina **dither**. Si bien el efecto es empeorar ligeramente la relación señal a ruido, desde el punto de vista auditivo transforma esta distorsión en un ruido aleatorio, que es mucho más aceptable, sobre todo en niveles tan bajos.

También se acostumbra a aplicar **dither** en los procesos de *recuantización*, es decir cuando, por ejemplo, se desea reducir la resolución de una señal grabada en 20 bits a 16 bits a fin de volcarla a un formato comercial como el compact disc. Si solamente se truncaran los datos de 20 bits eliminando los 4 bits menos significativos, se producirían inconvenientes similares al descrito. En ese caso, el ruido se genera digitalmente y se agrega antes de proceder al truncado.

Capítulo 16

Efectos I: Concepto y estructuras

16.1. Introducción

Los procesadores que hemos visto hasta el momento (por ejemplo los compresores) tenían la particularidad de que procesaban la totalidad de la señal que los atravesaba. Existe otro gran conjunto de procesadores denominados genéricamente **procesadores de efectos**, o simplemente **efectos**, que se utilizan mezclando la señal procesada (señal "húmeda") con parte de la señal sin procesar (señal "seca"). En términos generales los efectos agregan realismo, ambientación y espacialidad al sonido, o le dan extensión, o movimiento, o mayores posibilidades expresivas. También pueden alterar diversas características de los sonidos, como el contenido espectral, la frecuencia, o la envolvente.

16.2. Efectos en serie y en paralelo

En la terminología habitual se suele denominar **efectos en serie** o **efectos de inserción a los procesadores que procesan la totalidad de la señal** y **efectos en paralelo a los procesadores de efectos propiamente dichos**. En la Tabla 16.1 se incluye una lista de algunos efectos de ambos tipos, y en la Figura 16.1 se muestran las estructuras de conexionado de los mismos. Las conexiones generalmente disponibles en las consolas mezcladoras están motivadas por la existencia de dichas estructuras, y son de dos tipos: de **inserción** y **auxiliares**.

16.3. Conexiones de inserción

La conexión de **inserción** (en inglés, **insert**), consiste simplemente en un conector tipo **jack** con interrupción. En vacío (es decir sin conexión externa), la entrada queda directamente puenteadada hacia la salida, mientras que al insertar un **plug** la señal es derivada a través de una conexión de **envío** (en inglés, **send**) hacia el procesador, después del cual, ya procesada, reingresa a través de una conexión de **retorno** (en inglés, **return**).

82